ESSA Level 2 Evidence
Impact Research

i-Ready Learning
Magnetic Reading™

Impact of *Magnetic Reading* in Grades 3–5: Evidence in Iowa Schools in the 2021–2022 School Year

Madison A. Holzman, Ph.D. and Molly K. Duncan, Ph.D.

Curriculum Associates Research Report | February 2023

Overview

Curriculum Associates conducted a study examining the impact of *Magnetic Reading* on student reading achievement for students in Grades 3–5 from seven elementary schools.

To strengthen evidence that results are directly related to *Magnetic Reading* usage, *Magnetic Reading* and comparison students were matched on key characteristics including fall achievement and *i-Ready Personalized Instruction* (i.e., *i-Ready*) usage. Results from analyses conducted on the matched sample suggest that *Magnetic Reading* has a positive and significant effect on students' reading skills. Specifically, compared to similar peers:

- Students who used *Magnetic Reading* scored over eight points higher, on average, on their end-of-year reading assessment.
- More *Magnetic Reading* students progressed toward or met grade-level proficiency in reading.
- More *Magnetic Reading* students met their Typical Growth and Stretch Growth® targets in reading.

This study aligns with the Every Student Succeeds Act (ESSA)'s Moderate design principles and provides evidence that *Magnetic Reading* has a positive and significant impact on Grades 3–5 students' reading achievement and growth. Moreover, these findings support that *Magnetic Reading* may be a meaningful investment for supporting the acceleration of student learning.

Introduction

In fall 2021, Curriculum Associates launched *Magnetic Reading*, a print-based, teacher-led reading program designed to support students' grade-level comprehension skills development in Grades 3–5. The program's foundation comprises texts that are written to be high-interest, culturally relevant, and informational. The program's purposes are to support students' comprehension skills, build knowledge that fosters deeper learning and connection with the texts, and nurture a love of reading. For educators, *Magnetic Reading* offers a digital library of resources for whole class grade-level instruction and differentiated learning, as well as scaffolds and protocols with which all students can relate and participate. Students are intended to use the program daily for 30–45 minutes as one component of their longer and more robust reading block.

To understand how *Magnetic Reading* contributes to students' reading skills and knowledge, we conducted a research study designed to answer the question: What is the effect of *Magnetic Reading* on students' reading achievement? Reading achievement was measured by students' *i-Ready Diagnostic* for Reading (i.e., Diagnostic) assessment scores. Whereas it is useful to understand how the benefits of *Magnetic Reading* will influence students' scale scores, we know that scale scores do not tell the whole story. We know that readiness for grade-

level material and where students are in their readiness for grade level is also important—perhaps more important—than a change in scale score points alone. As such, we also evaluated the impact of *Magnetic Reading* from the lens of how the program influences students' journeys toward grade-level knowledge and skills. We also considered the question: Do more students meet their Typical Growth and Stretch Growth targets when using *Magnetic Reading*?

Specifically, this study answered these questions for students in small, rural, Title I-eligible schools in Iowa. All schools were in their first full year of *Magnetic Reading* implementation, when educators and students were learning and adjusting to the new program.

Methodology

All analyses were conducted in *R* version 4.1.3 (R Core Team, 2021).

Sample Identification

Students from seven schools in Iowa are represented in this study. The sample was identified by first selecting schools for inclusion, followed by selecting students for inclusion within those schools.

School Identification

Three schools in Iowa were identified as treatment schools. These schools implemented *Magnetic Reading* in Grades 3, 4, and 5 and reported using *Magnetic Reading* as their primary comprehension program for 30–45 minutes daily in their reading block for the full 2021–2022 academic year. According to the National Center for Education Statistics (NCES)'s most recent Common Core of Data (CCD; US Department of Education, 2022), treatment schools are small, rural, Title I eligible, and composed of 90% or more White students.

Due to limited sample size and school-level information, a statistical matching process to identify comparison schools was not feasible. Instead, NCES CCD data were used to identify comparison schools who were also small, rural, at least 90% White, Title I eligible, and served Grades 3, 4, and 5 in Iowa but did not use *Magnetic Reading* or *Ready® Reading.*[1] All identified treatment and comparison schools used *i-Ready Personalized Instruction*. However, *i-Ready* usage varied by school. Because *i-Ready* usage is known to have a positive effect on students'

[1]*Ready Reading* is a Curriculum Associates teacher-led reading program. It is the legacy product prior to *Magnetic Reading* and, like *Magnetic Reading*, is intended to serve as the primary comprehension program in students' daily reading blocks. Although *Ready Reading* and *Magnetic Reading* are distinct products, some content, features, and strategies are common between the programs. For this reason, schools using *Ready Reading* were removed from the sample of comparison schools.

achievement (Curriculum Associates, 2021a; Randel, Swain, Dvorak, & Prendez, 2020; Randel, Swain, Dvorak, Spratto, & Prendez, 2020; Swain et al., 2019), it was important to take *i-Ready* usage at the treatment and comparison schools into account to obtain an accurate estimate of the impact of *Magnetic Reading* on student achievement, rather than a mix of the impact *Magnetic Reading* and *i-Ready* usage. Accordingly, to be eligible for inclusion in the final school sample, comparison schools 1) were required to use *i-Ready* with at least 85% of their students in each Grades 3, 4, and 5 and 2) whose students used *i-Ready* for an average of at least 10 weeks in each Grades 3, 4, and 5. The final school sample consisted of three treatment schools and four comparison schools.

Student Identification

To be eligible for inclusion in the student sample, students in the identified schools were required to 1) complete a fall and spring *i-Ready Diagnostic* for Reading in school and 2) have used *i-Ready Personalized Instruction* for Reading. Some students in the sample had extreme *i-Ready* usage compared to the rest of the sample. Because such high *i-Ready* usage may represent a different instructional track than peers within the classroom, students who completed 87 or more lessons and spent 1,533 minutes or more on *i-Ready* (i.e., students who were in the top 10% of *i-Ready* usage) were removed.

Using the MatchIt Package in *R*, (Ho et al., 2011), propensity score matching was used to identify a sample of Grades 3–5 treatment and comparison students who were similar to one another within their grade on fall achievement and *i-Ready* usage, which were characteristics theorized to influence either end-of-year reading achievement (i.e., the outcome) or *Magnetic Reading* usage (i.e., the intervention to be tested). By accounting for characteristics theorized to influence the outcome or treatment usage, a less biased estimate of the effect of *Magnetic Reading* was obtained (Rosenbaum & Rubin, 1983).

To procure the matched sample, an iterative process that started with nearest neighbor matching without a caliper was used, with the addition of various calipers until covariate balance was maximized and loss of *Magnetic Reading* students was minimized (Harris & Horst, 2016). Covariates were considered to have adequate balance if the standardized mean difference between *Magnetic Reading* and comparison groups after matching was < |.25| (What Works Clearinghouse, 2022). The final matching model used nearest neighbor matching with a .20 caliper and exact match on grade to predict *Magnetic Reading* assignment from fall Diagnostic for Reading score, total time spent on *i-Ready*, total *i-Ready* lessons completed, and number of weeks using *i-Ready* between fall and spring Diagnostics. See Table 1 for sample balance before and after matching.

Table 1. Covariate Balance before and after Nearest Neighbor Matching with a .20 Caliper

Before Matching			
	Magnetic Reading	**Comparison**	**|Standardized Mean Difference|**
N	223	339	—
Fall Diagnostic Score	522.42	522.49	.00
Total Time Spent on *i-Ready*	792.07	915.89	.35
Total *i-Ready* Lesson Count	43.66	44.53	.04
Number of Weeks Spent Using *i-Ready*	24.53	27.47	.41
After Matching			
	Magnetic Reading	**Comparison**	**|Standardized Mean Difference|**
n	179	179	—
Fall Diagnostic Score	526.16	515.57	.19
Total Time Spent on *i-Ready*	843.78	885.77	.12
Total *i-Ready* Lesson Count	45.10	48.64	.17
Number of Weeks Spent Using *i-Ready*	26.11	26.93	.12

Impact Model

The first research question sought to evaluate the effect of *Magnetic Reading* on students' reading achievement. The impact of *Magnetic Reading* on spring reading achievement was evaluated via a linear regression analysis on the matched Grades 3–5 sample. Using the stats package in *R* (R Core Team, 2021), a series of nested and non-nested models that included fall Diagnostic for Reading scores and various *i-Ready* metrics were compared to one another to determine the best combination of predictors to evaluate the impact of *Magnetic Reading* on spring Diagnostic scores. The final model predicted students' spring Diagnostic scores from their fall Diagnostic score, total time spent on *i-Ready*, their grade, an interaction of their grade and fall

Diagnostic score, and *Magnetic Reading* usage. To accommodate interpretation of the eventual model coefficients, fall Diagnostic score and total time spent on *i-Ready* were centered. Fall Diagnostic score was centered at 500 because 500 is a plausible score for each Grade 3, 4, or 5. Total time spent on *i-Ready* was centered at 540 minutes because 540 minutes represents our suggested usage guidance of at least 30 minutes of *i-Ready* usage per week for at least 18 weeks in the school year. The final model explained 75.5% of variability in spring Diagnostic scores. Assumption checks were conducted, and no major violations were identified.

Given the data were clustered within schools, cluster-robust standard errors were employed. Although hierarchical linear modeling (HLM) is the more common technique used to account for clustered data in education, cluster-robust standard errors were employed as an alternative to HLM because the sample size within schools was small, there were few schools in the sample, and the school effect was not of substantive interest for our research questions (McNeish et al., 2017). Cluster-robust standard errors were obtained using the miceadds package in *R* (Robitzsch & Grund, 2022).

After estimating the regression coefficients, Cohen's *d* was calculated as a standardized effect size for the impact of *Magnetic Reading*. Cohen's *d* was standardized to the pooled standard deviation of the raw spring Diagnostic score between the groups. To contextualize the range of the plausible effect of *Magnetic Reading* on spring achievement, 95% confidence intervals were calculated around the *Magnetic Reading* impact estimate using the cluster-robust standard errors.

Grade-Level Placement Changes

To contextualize the effect of *Magnetic Reading* on students' reading achievement, placement changes from fall to spring were evaluated using the matched sample. At fall and spring, students' placements were combined into three placements relative to their grade level: Two or More Grade Levels Below, One Grade Level Below, and Early On Grade Level or Above. Placements were defined by Curriculum Associates' "Standard View"[2] at both fall and spring. The fall relative grade-level placement counts and percentages for the matched sample are in Table 2. Students' transitions in relative grade-level placement from fall to spring were then analyzed descriptively.

[2]Per the "Standard View," students who placed Early On Grade, Mid On Grade, or above their chronological grade were considered to be Early On Grade Level or Above.

Table 2. Fall Relative Grade-Level Placement Counts and Percentages for the Final Matched Sample

	Two or More Grade Levels Below	One Grade Level Below	Early On Grade Level or Above	Total
Magnetic Reading	55 (31%)	71 (40%)	53 (30%)	179
Comparison	78 (44%)	52 (29%)	49 (27%)	179
Total	133	123	102	358

Growth Targets

The second research question sought to evaluate whether more students met their Typical Growth and Stretch Growth targets if they used *Magnetic Reading*. Curriculum Associates' Typical Growth and Stretch Growth are used to set goals for individual students or groups of students. Growth targets are determined based on students' grades and fall placements on the Diagnostic for a given subject. Typical Growth is the fall-to-spring growth expected for the average student on their Diagnostic. For many students, meeting Typical Growth will likely help them maintain their grade-level placement. This means that for students who are below grade level, meeting Typical Growth is insufficient for achieving grade-level proficiency. Stretch Growth is an ambitious yet attainable goal representing the fall-to-spring growth a student should aim for to be on a path toward grade-level proficiency. Students' Typical Growth and Stretch Growth goals are provided to educators in *i-Ready* reports and are suggested focus areas when setting goals with individual students or groups of students.

Students were considered to have met their Typical Growth or Stretch Growth targets if their spring Diagnostic score was at or above their respective growth target. The percentage of students in the matched sample who met their Typical Growth or Stretch Growth targets was calculated for the *Magnetic Reading* and comparison groups. Using the gmodels package (Warnes et al., 2022) in *R*, a chi-square test of independence was used to evaluate whether there was an association between *Magnetic Reading* usage and meeting Typical Growth or Stretch Growth targets. Phi (ϕ) is provided as the effect size. Phi ranges from 0–1, with values closer to 1 representing a stronger relationship between *Magnetic Reading* usage and meeting the respective growth targets.

To further interpret the influence of *Magnetic Reading* on Typical Growth or Stretch Growth goal attainment, the relative risk associated with *Magnetic Reading* usage and meeting Typical Growth or Stretch Growth was calculated. The relative risk represents how much more likely students are to meet their respective growth target if they use *Magnetic Reading*. Whereas "relative risk" is the language used in the literature, in this case the language of relative risk is non-intuitive because the goal is for students to achieve their growth targets. Instead, we may interpret findings as the "relative benefit" of *Magnetic Reading* for meeting Typical Growth or Stretch Growth targets.

Findings

What is the effect of *Magnetic Reading* on students' reading achievement?

In small, rural schools in Iowa, Grades 3–5 students using *Magnetic Reading* scored, on average, more than eight points higher on their spring Diagnostic compared to similar students who did not use *Magnetic Reading* (see Table 3). Considering we expect the impact of *Magnetic Reading* to vary by student, the confidence interval of the program's effect is a useful metric to consider. The confidence interval represents the plausible range of scale score points *Magnetic Reading* students should expect to score higher than their similar peers who do not use *Magnetic Reading*. In this sample and context, the impact of *Magnetic Reading* may range from a nearly three-point to 13.5-point advantage on students' spring Diagnostic. This finding is after taking into account students' fall Diagnostic scores, total time spent on *i-Ready* in the academic year, and grade. Notably, because *Magnetic Reading* and comparison students were matched on key *i-Ready* metrics such as total time spent using *i-Ready*, total lessons completed, and number of weeks using *i-Ready* between fall and spring Diagnostics, the estimated effect of using *Magnetic Reading* represents the impact of *Magnetic Reading* above and beyond *i-Ready* usage.

The Cohen's *d*, or standardized effect size, associated with the 8.25 scale score benefit of *Magnetic Reading* is .18 standard deviations. Based on Lipsey and colleagues' (2012) and Kraft's (2020) research on the effects of educational interventions, the standardized effect size of *Magnetic Reading* is considered moderate and practically meaningful. See Appendix A for the full table of results from the impact model.

Table 3. Impact of *Magnetic Reading* on Spring Reading Achievement

	Scale Score Difference	Cluster-Robust Standard Error	Confidence Interval	Standardized Effect Size
Magnetic Reading	8.25**	2.70	2.96–13.54	.18

Note. Scale score difference represents the mean difference between groups, adjusted for fall Diagnostic score, total time spent on *i-Ready*, grade, and the interaction between grade and fall Diagnostic score.
Note. **p <= .01.
Note. Confidence interval is calculated using $\alpha = .05$.

To further contextualize what the difference in scores between groups means for students, placement change from fall to spring was also examined. Notably, nearly three-fourths of *Magnetic Reading* students who placed One Grade Level Below in the fall placed on or above grade level in the spring. This is compared to 60% of comparison students who ended on or above grade level when starting the year one grade level below. Additionally, half of the *Magnetic Reading* students who placed Two or More Grade Levels Below in the fall ended the year only one grade level below. This is compared to only one-third of comparison students with the same change in placement. Although sample sizes are small and percentages should be interpreted with caution, these results are promising and suggest *Magnetic Reading* may help Grades 3–5 students approach or

reach grade-level expectations. See Table 4 for the placement transition table. Percentages represent the percentage of students in either the *Magnetic Reading* or comparison group in each placement category.

Table 4. Placement Transition from Fall to Spring by Group and Grade-Level Placement

Magnetic Reading					
Spring Diagnostic Placement					
		n	**Two or More Grade Levels Below**	**One Grade Level Below**	**Early On Grade Level or Above**
Fall Diagnostic Placement	**Two or More Grade Levels Below**	55	16 (29.09%)	28 (50.91%)	11 (20.00%)
	One Grade Level Below	71	2 (2.82%)	17 (23.94%)	52 (73.24%)
	Early on Grade Level or Above	53	0 (0%)	3 (5.66%)	50 (94.34%)
Comparison					
Spring Diagnostic Placement					
		n	**Two or More Grade Levels Below**	**One Grade Level Below**	**Early On Grade Level or Above**
Fall Diagnostic Placement	**Two or More Grade Levels Below**	78	38 (48.72%)	26 (33.33%)	14 (17.95%)
	One Grade Level Below	52	1 (1.92%)	20 (38.46%)	31 (59.62%)
	Early on Grade Level or Above	49	0 (0%)	4 (8.16%)	45 (91.84%)

Note. Percentages represent the percentage of students within a *Magnetic Reading* or comparison group who began the year in a given placement and ended in a given placement. Sample interpretation: Of the *Magnetic Reading* students who started two or more grade levels below in the fall, 20% experienced growth that resulted in on-grade level placement in the spring.

Do more students meet their Typical Growth and Stretch Growth targets when using *Magnetic Reading*?

In this sample, significantly more *Magnetic Reading* learners met their Typical Growth and Stretch Growth targets than similar comparison students (see Table 5). 77% of *Magnetic Reading* students met their Typical Growth goals, compared to only 65% of comparison students. 44% of *Magnetic Reading* students met their Stretch Growth goals, compared to only 27% of comparison students. In terms of students, this means that of 179 students in each group, 21 more *Magnetic Reading* students met their Typical Growth targets and 30 more *Magnetic Reading* students met their Stretch Growth targets compared to their non-*Magnetic Reading* peers.

Furthermore, students using *Magnetic Reading* were nearly 1.2 times more likely to meet their Typical Growth targets and 1.6 times more likely to meet their Stretch Growth targets than their similar peers. These results suggest that *Magnetic Reading* may support students in meeting their growth targets as well as support students' growth toward grade-level proficiency. See Table 5 for the counts and percentages of students who met their Typical Growth and Stretch Growth targets.

Table 5. Percentage of Students who Met Typical and Stretch Growth Targets by Group

	Typical Growth		Stretch Growth	
	Did Not Meet Typical Growth Target	Met Typical Growth Target	Did Not Meet Stretch Growth Target	Met Stretch Growth Target
Magnetic Reading	41 (22.91%)	138 (77.09%)	100 (55.87%)	79 (44.13%)
Comparison	62 (34.64%)	117 (65.36%)	130 (72.63%)	49 (27.37%)
Percentage Point Difference	-11.73	11.73	-16.76	16.76

Typical Growth [$\chi 2(1, N = 358) = 6.01, p = .01, \phi = 0.13$]

Stretch Growth [$\chi 2(1, N = 358) = 10.94, p < .001, \phi = 0.18$]

Note. Percentage point difference was calculated by subtracting the comparison percentage from the *Magnetic Reading* percentage. A positive value indicates a greater percentage of *Magnetic Reading* students met their growth targets.

Discussion

Acceleration toward Grade-Level Expectations

The findings from this study are promising and reflect that *Magnetic Reading* has a positive impact on students' reading achievement. In particular, findings reflect that *Magnetic Reading* supports students in meeting their

growth target goals. These findings are particularly relevant as educators, students, and their caregivers continue to navigate teaching and learning after disruptions related to the COVID-19 pandemic. Data for this study were from the 2021–2022 school year, just one year after teaching and learning disruptions due to the COVID-19 pandemic in the 2020–2021 school year. We know that the teaching and learning disruptions—as well as the mental, emotional, and financial toll of the pandemic—negatively impacted students' educational experiences and resulted in unfinished learning that is unprecedented in our modern education system. Moreover, we know the negative effects of the pandemic on education were more severe for marginalized groups of students, namely students in schools who serve primarily Latino, Black, and economically disadvantaged students (Curriculum Associates, 2021b) as well as students who are below grade level (Dawson, 2022; Lewis et al., 2022). Consequently, it is more important than ever to ensure learners have effective instructional materials that not only support their knowledge and skills development, but also set them on the path toward grade-level proficiency and support their acceleration of gaining grade-level skills and knowledge.

Curriculum Associates' Stretch Growth makes tangible for educators, students, and their caregivers what goals students must meet to achieve grade-level skills and knowledge. Stretch Growth targets are ambitious, and meeting Stretch Growth puts students on the path toward grade-level proficiency. In fact, for Grades 3–5 students who meet Stretch Growth two years in a row, the vast majority reach grade-level proficiency at the end of the second year (Rome & Daisher, 2022). While Stretch Growth targets are attainable with the appropriate scaffolds and support, few students meet Stretch Growth in one school year, let alone in two consecutive years. As such, we need to understand how to better support students in meeting their Stretch Growth targets.

This study demonstrates that *Magnetic Reading* may help more students meet their Stretch Growth targets and move toward or achieve grade-level proficiency. In the 2017–2018 and 2018–2019 school years, 29% to 35% of students in Grades 3–5 met their reading Stretch Growth targets in a single year (Rome & Daisher, 2022). In this study, 44% of students who used *Magnetic Reading* met their Stretch Growth targets. Their peers who did not use *Magnetic Reading* met Stretch Growth at levels just lower than those observed historically. Moreover, an additional 30 students met Stretch Growth if they used *Magnetic Reading*. Thirty students is nearly two classrooms of students who are now on track toward grade-level proficiency and who may not have been without the opportunity to use *Magnetic Reading*. *Magnetic Reading* provides an opportunity for educators to support students' Stretch Growth attainment, and thus an opportunity for students to achieve the educational outcomes to which they aspire.

Limitations and Future Directions

This research provides encouraging evidence for *Magnetic Reading* and the academic benefits of the program for students in Grades 3–5. Yet, it is important to acknowledge that this study used a small and specific sample relative to the national population of Grades 3–5 students in the United States. Students in this sample were primarily White and attended small, rural, Title I-eligible schools in Iowa. Future research will focus on broadening our evaluation of *Magnetic Reading* by including students and schools with a wider range of diverse characteristics and identities.

We know that implementations of *Magnetic Reading* vary by schools and classrooms. Moreover, we know the other curricular materials used in addition to *Magnetic Reading* vary by schools and classrooms. In this study, we did not have information regarding classrooms' implementation of *Magnetic Reading*, nor about their other programs or materials they used alongside of *Magnetic Reading*. Because we could not take classroom-level information into account, it is possible the effects of *Magnetic Reading* presented in this paper could be misattributed to unknown and unmeasured factors in the classroom.

Conclusion

This study answered two questions related to the efficacy of *Magnetic Reading*, a print-based reading comprehension program for students in Grades 3–5:

1. What is the effect of *Magnetic Reading* on Grade 3–5 students' reading achievement?
2. Do more students meet their Typical Growth and Stretch Growth targets when using *Magnetic Reading*?

The results presented in this paper reflect that *Magnetic Reading* has a positive and significant impact on Grades 3–5 students' reading achievement and growth. For students in small, rural, Title I-eligible schools in Iowa, *Magnetic Reading* had a positive influence on their *i-Ready Diagnostic* for Reading achievement and their growth toward meeting grade-level expectations. Students who used *Magnetic Reading* scored more than eight points higher on their spring Diagnostic, more *Magnetic Reading* students met or progressed toward grade-level expectations, and more *Magnetic Reading* students met their Typical Growth and Stretch Growth targets in reading compared to similar peers who did not use *Magnetic Reading*. Notably, these results represent the impact of *Magnetic Reading* above and beyond the benefits students likely already experience from *i-Ready* usage.

Although this was a small study, findings are promising and suggest that *Magnetic Reading* supports students in achieving their educational goals. Future research will commit to broadening our understanding of the impact of *Magnetic Reading* for all students across educational contexts, school environments, and student identities.

References

Curriculum Associates. (2021a). *Impact of* i-Ready Personalized Instruction *time and lesson pass rates on student learning gains*. Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-instruction-dosage-study-2021.pdf

Curriculum Associates. (2021b). *Understanding student learning: Insights from fall 2021*. Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-understanding-student-learning-paper-fall-results-2021.pdf

Dawson, M. (2022). *Student growth during COVID-19: Grade-level readiness matters.* Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/student-growth-during-covid-19.pdf

Harris, H. & Horst, S. J. (2016). A brief guide to decisions at each step of the propensity score matching process. *Practical Assessment, Research, and Evaluation, 21*(4). https://doi.org/10.7275/yq7r-4820

Ho, D., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software, 42*(8), 1–28. https://doi.org/10.18637/jss.v042.i08

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4)*,* 241–253. https://journals.sagepub.com/doi/full/10.3102/0013189X20912798

Lewis, K., Kuhfeld, M., Langi, M., Peters, S., Fahle, E. (2022). *The widening achievement divide during COVID-19.* NWEA Center for School and Student Progress. https://www.nwea.org/uploads/2022/11/CSSP-Brief_Widening-achievement-divide-COVID-19.pdf

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. Institute of Education Sciences, National Center for Special Education Research. https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*(1)*,* 114–140. https://psycnet.apa.org/doiLanding?doi=10.1037%2Fmet0000078

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Randel, B., Swain, M., Dvorak, R. N., & Prendez, J. Y. (2020). *Effectiveness of* i-Ready Instruction *in elementary reading for traditionally disadvantaged groups.* Human Resources Research Organization.

Randel, B., Swain, M., Dvorak, R. N., Spratto, E., & Prendez, J. Y. (2020). *Impact evaluation of reading* i-Ready *for striving learners using 2018–19 data: Final report.* (Report No. 053). Human Resources Research Organization. https://files.eric.ed.gov/fulltext/ED610441.pdf

Robitzsch, A., & Grund, S. (2022). *miceadds: Some additional multiple imputation functions, especially for 'mice.'* R package version 3.15-21. https://CRAN.R-project.org/package=miceadds

Rome, L. & Daisher, T. (2022). i-Ready *Stretch Growth as a path toward proficiency.* Curriculum Associates. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/ireadytwoyearstretchgrowthresearchtechnicalreview20220913-1.pdf

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1)*,* 41–55. https://doi.org/10.1093/biomet/70.1.41

Swain, M., Randel, B., & Dvorak, R. N. (2019). *Impact evaluation of reading* i-Ready *instruction for elementary grades using 2018–19 data*: *Final report.* (Report No. 107). Human Resources Research Organization. https://files.eric.ed.gov/fulltext/ED604746.pdf

US Department of Education. (2022). Institute of Education Sciences, National Center for Education Statistics. https://nces.ed.gov/ccd/ccddata.asp

Warnes, G. R., Bolker, B., Lumley, T., & Johnson, R. C. (2022). *Gmodels: Various R programming tools for model fitting.* R package version 2.18.1.1. https://CRAN.R-project.org/package=gmodels

What Works Clearinghouse. (2022). *Procedures and standards handbook, version 5.0.* US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/Final_WWC-HandbookVer5_0-0-508.pdf

Appendix A.

Table 6. Regression Estimates

Variable	Category	Unstandardized Estimate	Cluster-Robust Standard Error	*p*
Intercept		524.87	4.35	< .001
Fall Diagnostic		.70	.036	< .001
***Magnetic Reading* Usage (Reference = Comparison Group)**		8.25	2.70	< .05
Total Time Spent Using *i-Ready*		.01	.01	.01
Grade (Reference = Grade 3)	**Grade 4**	9.42	6.15	.13
	Grade 5	1.74	5.71	.76
Fall Diagnostic Score by Grade Interaction (Reference = Grade 3)	**Grade 4**	-.013	.05	.79
	Grade 5	.07	.05	.18

Note. Fall Diagnostic for Reading score was centered at 500, and total time spent on *i-Ready* was centered at 540 minutes.
Note. Unstandardized estimates represent the average difference between groups, adjusted for the other covariates in the model.